Runtime randomization and perturbation for virtual machines.

JAVIER CABRERA ARTEAGA

Licentiate Thesis in [Research Subject - as it is in your ISP]
School of Information and Communication Technology
KTH Royal Institute of Technology
Stockholm, Sweden [2022]

KTH School of Information and
Communication Technology
SE-164 40 Kista
SWEDEN

TRITA-ICT XXXX:XX
ISBN XXX-XX-XXXX-XXX-X

Akademisk avhandling som med tillstånd av Kungl Tekniska högskolan framlägges till offentlig granskning för avläggande av licentiatexamen i [ämne/subject] [veckodag/weekday] den [dag/day] [månad/month] [år/2022] klockan [tid/time] i [sal/hall], Electrum, Kungl Tekniska högskolan, Kistagången 16, Kista.

© Javier Cabrera Arteaga, [month] [2022]

Tryck: Universitetsservice US AB

Abstract

Write your abstract here...
Keywords: Keyword1, keyword2, ...

Sammanfattning

Write your Swedish summary (popular description) here...

Keywords: Keyword1, keyword2, ...

Acknowledgements

Write your professional acknowledgements here...

Acknowledgements are used to thank all persons who have helped in carrying out the research and to the research organizations/institutions and/or companies for funding the research.

Name Surname,
Place, Date

Contents

Contents **vi**

List of Figures **viii**

List of Tables **ix**

List of Acronyms **xi**

1 Introduction **1**
1.1 Motivation 1
1.1.1 Why variants ? 1
1.1.2 Research questions 1
1.2 Contributions 1

2 Background & State of the art **3**
2.1 WebAssembly 3
2.2 Diversification and Superoptimization for Superdiversification 5
2.3 Security applications for diversification 7
2.4 Conclusions 7

3 Methodology **9**
3.1 RQ1. To what extent can we artifically generate program variants for WebAssembly? 10
3.2 RQ2. To what extent are the generated variants dynamically different? 13
3.3 RQ3. To what extent do the artificial variants exhibit different execution times on Edge-Cloud platforms? 16

4 Results **19**
4.1 RQ1. To what extent can we artifically generate program variants for WebAssembly? 19
4.2 RQ2. To what extent are the generated variants dynamically different? 22
4.3 RQ3. To what extent do the artificial variants exhibit different execution times on Edge-Cloud platforms? 25

Bibliography **29**

List of Figures

3.1 The program variants generation for RQ1. 12
3.2 Dynamic analysis for RQ2. 14
3.3 Multivariant binary creation and workflow for RQ3 answering. 16

4.1 Pairwise comparison of programs' population traces in logarithmic scale. Each vertical group of blue dots represents a programs' population. Each dot represents a comparison between two program execution traces according to Metric 5. 23
4.2 Execution time distributions for `Hilber_curve` program and its variants. Baseline execution time mean is highlighted with the magenta horizontal line. 24
4.3 Execution time distributions. Each subplot represents the quantile-quantile plot of the two distributions, original and multivariant binary. 27

List of Tables

3.1 Corpora description. The table is composed by the name of the corpus, the number of modules, the number of functions, the lines of code range and the location of the corpus. 11

4.1 General program's populations statistics. The table is composed by the name of the corpus, the number of functions, the number of succesfully diversified functions, the cumulative number of generated variants and the cumulative number of unique variants. 20

List of Acronyms

Wasm	WebAssembly
DTW	Dynamic Time Warping